数理解析研究所講究録
第 887 巻 1994 年 97-106

無限次元空間上の測度と情報理論

Dedicated to Prof. Yasuo Yamasaki on his sixtieth birthday

山口大・工　柳　研二郎 (Kenjiro Yanagi)

1　はじめに

無限次元空間上の測度の応用の1つとして1948年シャノンによって創設された情報理論がある。情報理論は非常に広領域の分野を含む学際領域に位置する新しい分野を形成し、現在では解析学、確率論のみならず符号理論に関しては代数幾何学の手法を取り込んでめざましい成果を挙げている。一方通信路の数学的構成は一般の確率測度空間を用いてある程度完成しているといえるが (国澤 - 梅垣 [8]) 細かい議論に関しては未解決問題が多数残っている。この論文では入力空間および出力空間が無限次元空間の場合を想定し特にガウス型の通信路の構成を行う。なぜ無限次元空間かといえば、入力信号として連続時間確率過程を考えれば当然そのパス空間は無限次元空間となる。その上の確率測度が入力信号の分布の役割を演ずる。
まず第1章では無限次元空間の中でも最もポピュラーなバナッハ空間上の確率測度についてのよく知られた結果を述べる。これを用いてフィードバックを持たないガウス型通信路を定義し相互情報量や容量について言及する。
第2章ではフィードバックをもつガウス型通信路を扱うがフィードバックをもつ場合には測度による表現が不可能で時間パラメータが直接表に表れる確率過程による表現が必要になってくる。ここではまず離散時間の場合を扱い容量についての2、3の結果を与える。
最後に第3章で連続時間の場合を扱う。

2　バナッハ空間上の確率測度

X を実可分バナッハ (Banach) 空間, X^* をその共役空間とする. $\mathcal{B}(X)$ を X の Borel σ- 集合体とする. X^* の有限次元部分空間 F に対して F に基づいた柱状集合 (cylinder set) C は次のように定義される.

$$C = \{x \in X; (< x, f_1 >, < x, f_2 >, \dots, < x, f_n >) \in D\}$$

ただし $n \geq 1$, $\{f_1, f_2, \ldots, f_n\} \subset F$, $D \in \mathcal{B}(\mathbb{R})$ である. F に基づいた柱状集合全体を $\mathcal{C}_F$ とし,

$$\mathcal{C}(X, X^*) = \bigcup\{\mathcal{C}_F; F \text{ は } X^* \text{ の有限次元部分空間 }\}$$

とすると $\mathcal{C}(X, X^*)$ は集合体となる. $\hat{\mathcal{C}}(X, X^*)$ を $\mathcal{C}(X, X^*)$ によって生成される $\sigma-$集合体とすると $\hat{\mathcal{C}}(X, X^*) = \mathcal{B}(X)$ が成り立つ. μ を $\int_X ||x||^2 d\mu(x) < \infty$ を満たす $(X, \mathcal{B}(X))$ 上の確率測度とすると, 次のような vector $m \in X$ と operator $R : X^* \to X$ が存在することがわかる. つまり任意の $x^* \in X^*, y^* \in Y^*$ に対して

$$< m, x^* >= \int_X < x, x^* > d\mu(x),$$

$$< Rx^*, y^* >= \int_X < x - m, x^* >< x - m, y^* > d\mu(x)$$

である. この m を μ の平均ベクトル (mean vector) という. R は有界線型作用素であり, μ の共分散作用素 (covariance operator) という. さらにこの R は対称 (symmetric) であり, つまり

$$\text{任意の } x^*, y^* \in X^* \text{ に対して } < Rx^*, y^* >=< Ry^*, x^* >$$

また正定値 (positive) でもある. つまり

$$\text{任意の } x^* \in X^* \text{ に対して } < Rx^*, x^* >\geq 0$$

任意の $f \in X^*$ に対して $\mu_f = \mu \circ f^{-1}$ が $\mathbb{R}$ 上のガウス測度となるとき μ を $(X, \mathcal{B}(X))$ 上のガウス測度という. ガウス測度 μ の特性函数 $\hat{\mu}(f)$ は次のように表される. 任意の $f \in X^*$ に対して

$$\hat{\mu}(f) = exp\{i < m, f > -\frac{1}{2} < Rf, f >\} \tag{1}$$

ただし $m \in X$ は μ の平均ベクトル, $R : X^* \to X$ は μ の共分散作用素である. 逆に $(X, \mathcal{B}(X))$ 上の確率測度 μ の特性函数が (1) の形をしていれば μ はガウス測度となり $m \in X$ はその平均ベクトル, $R : X^* \to X$ はその共分散作用素となっている. したがって $\mu = [m, R]$ と書いて μ は平均ベクトル m, 共分散作用素 R をもつ $(X, \mathcal{B}(X))$ 上のガウス測度を表すことにする.

任意に symmetric positive operator $R : X^* \to X$ が与えられると X のヒルベルト (Hilbert) 部分空間 H と H から X への連続な埋め込み j が存在して $R = jj^*$ となる. この辺の詳しい議論については Vakhania-Tarieladza-Chobanyan [12] を見よ. この H を R の再生核ヒルベルト空間 (reproducing kernel Hilbert space) という. なぜこのような名前がつけられることが可能かというと次の理由のためである. ヒルベルト空間 $\mathcal{H}$ と X^* から $\mathcal{H}$ への有界線型作用素 A が存在して $A^*A = R$ かつ $A(X^*)$ は $\mathcal{H}$ で稠密である. ここで $\mathcal{H}_R = A^*(\mathcal{H})$ とする. また $k_R(x^*, y^*) =< Rx^*, y^* >$ と定義

すると k_R は $X^* \times X^*$ 上の正定値函数となる. この k_R によってつくられる再生核ヒルベルト空間を $\mathcal{H}(k_R)$ とすると $\mathcal{H}(k_R) \cong \mathcal{H}_{\mathcal{R}}$ である.
次に相互情報量を定義するために2つの実可分バナッハ空間 X, Y を用意する. μ_X, μ_Y をそれぞれ $(X, \mathcal{B}(X)), (Y, \mathcal{B}(Y))$ 上の確率測度, μ_{XY} を μ_X, μ_Y をそれぞれ周辺分布にもつような $(X \times Y, \mathcal{B}(X) \times \mathcal{B}(Y))$ 上の結合確率測度とする. つまり

$$\text{任意の } A \in \mathcal{B}(X) \text{ に対して } \mu_X(A) = \mu_{XY}(A \times Y)$$

$$\text{任意の } B \in \mathcal{B}(Y) \text{ に対して } \mu_Y(B) = \mu_{XY}(X \times B)$$

が満たされる. さらに

$$\int_X ||x||^2 d\mu_X(x) < \infty, \quad \int_Y ||y||^2 d\mu_Y(y) < \infty$$

を仮定すると $m = (m_1, m_2) \in X \times Y$ が存在して次を満たす.
任意の $(x^*, y^*) \in X^* \times Y^*$ に対して

$$< (m_1, m_2), (x^*, y^*) >= \int_{X \times Y} < (x, y), (x^*, y^*) > d\mu_{XY}(x, y)$$

ただし m_1, m_2 はそれぞれ μ_X, μ_Y の平均ベクトルである. また

$$\mathcal{R} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} : X^* \times Y^* \to X \times Y$$

が存在して次を満たす.
任意の $(x^*, y^*), (z^*, w^*) \in X^* \times Y^*$ に対して

$$\begin{aligned} &< \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \begin{pmatrix} x^* \\ y^* \end{pmatrix}, \begin{pmatrix} z^* \\ w^* \end{pmatrix} > \\ = &\int_{X \times Y} < (x, y) - m, (x^*, y^*) >< (x, y) - m, (z^*, w^*) > d\mu_{XY}(x, y) \end{aligned}$$

ただし $R_{11} : X^* \to X$ は μ_X の共分散作用素, $R_{22} : Y^* \to Y$ は μ_Y の共分散作用素である. $R_{12} = R_{21}^* : Y^* \to X$ は次によって定義される.
任意の $(x^*, y^*) \in Y^* \times X^*$ に対して

$$< R_{12} y^*, x^* >= \int_{X \times Y} < x - m_1, x^* >< y - m_2, y^* > d\mu_{XY}(x, y)$$

この R_{12} は μ_{XY} の交差共分散作用素 (cross covariance operator) という.
$\mu_{XY} = \left[(0,0) \quad \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \right]$ とすると $\mu_X = [0, R_X]$, $\mu_Y = [0, R_Y]$ となる. また R_X の再生核ヒルベルト空間 $H_X \subset X$ と R_Y の再生核ヒルベルト空間 $H_Y \subset Y$, さら

に H_X から X への連続な埋め込み j_X と H_Y から Y への連続な埋め込み j_Y が存在し $R_X = j_X j_X^*$, $R_Y = j_Y j_Y^*$ となる.
ここでさらに再生核ヒルベルト空間 H_X が X で稠密, H_Y が Y で稠密と仮定すると有界線型作用素 $V_{XY} : H_Y \to H_X$ が存在して

$$R_{XY} = j_X V_{XY} j_Y, \quad \|V_{XY}\| \leq 1$$

を満たすようにできる. したがって次のような定理にまとめられる.

Theorem 1 $\mu_{XY} \sim \mu_X \otimes \mu_Y$ であるための必要十分条件は V_{XY} がヒルベルト・シュミット型 *(Hilbert-Schmidt type)* で $\|V_{XY}\| < 1$ である.

情報理論で特に重要な役割を演ずる μ_{XY} の相互情報量 $I(\mu_{XY})$ は次のように定義される.

$$\mathcal{F} = \{(\{A_j\}, \{B_j\}); \{A_j\} \text{ は } \mu_X(A_j) > 0 \text{ となる } X \text{ の有限可測分割}, \{B_j\} \text{ は } \mu_Y(B_j) > 0 \text{ となる } Y \text{ の有限可測分割 } \}$$

とすると

$$I(\mu_{XY}) = \sup \sum_{i,j} \mu_{XY}(A_i \times B_j) \log \frac{\mu_{XY}(A_i \times B_j)}{\mu_X(A_i)\mu_Y(B_j)}$$

である. ただし上限はすべての $(\{A_i\}, \{B_j\}) \in \mathcal{F}$ についてとる.
この相互情報量は次のように表される.
$\mu_{XY} \ll \mu_X \otimes \mu_Y$ のとき

$$I(\mu_{XY}) = \int_{X \times Y} \log \frac{d\mu_{XY}}{d\mu_X \otimes \mu_Y}(x, y) d\mu_{XY}(x, y)$$

その他のとき $I(\mu_{XY}) = \infty$ とする.
以上より次の定理が成り立つ.

Theorem 2 $\mu_{XY} \sim \mu_X \otimes \mu_Y$ のとき $I(\mu_{XY}) < \infty$ で

$$I(\mu_{XY}) = -\frac{1}{2} \sum_{n=1}^{\infty} \log(1 - \gamma_n)$$

と表される. ただし $\{\gamma_n\}$ は $V_{XY}^* V_{XY}$ の固有値である.

いよいよフィードバックをもたないガウス型通信路を定義することができる.
X を入力空間を表す実可分バナッハ空間, Y を出力空間を表す実可分バナッハ空間とする. $\lambda : X \times \mathcal{B}(Y) \to [0, 1]$ は次の (1), (2) を満たすとする.

(1) 任意の $x \in X$ に対して $\lambda(x, \cdot) = \lambda_x$ は $(Y, \mathcal{B}(Y))$ 上のガウス測度である

(2) 任意の $B \in \mathcal{B}(Y)$ に対して $\lambda(\cdot, B)$ は $(X, \mathcal{B}(X))$ 上のボレル可測函数である.

このとき3つ組 $[X, \lambda, Y]$ をガウス型通信路という.
入力情報源 μ_X を与えるとそれに対応して出力情報源 μ_Y 及び複合情報源 μ_{XY} がそれぞれ次のように定義される.
任意の $B \in \mathcal{B}(Y)$ に対して

$$\mu_Y(B) = \int_X \lambda(x, B) d\mu_X(x),$$

任意の $C \in \mathcal{B}(X) \times \mathcal{B}(Y)$ に対して

$$\mu_{XY}(C) = \int_X \lambda(x, C_x) d\mu_X(x)$$

ただし $C_x = \{y \in Y; (x, y) \in C\}$ である.
通信路の容量とはある制約条件を満たす入力情報源 μ_X に対して相互情報量 $I(\mu_{XY})$ の上限である. 容量の重要性はシャノンの第2符号化定理から保証されているのでただ単なる数学の遊びではないことに注意しておく.
簡単のため $X = Y$ で $\lambda(x, B) = \mu_Z(B - x)$, $\mu_Z = [0, R_Z]$ とする. つまり入力空間と出力空間は同じで雑音にあたるガウス測度 μ_Z が加法的に加わる通信路である. 制限として $\int_X \|x\|_Z^2 d\mu_X(x) \le P$ を与えると, その容量は $\frac{P}{2}$ であることが示される.

3 フィードバックをもつ離散時間ガウス型通信路

次のようなフィードバックをもつ離散時間ガウス型通信路を考える。

$$Y_n = S_n + Z_n, \quad n = 1, 2, \ldots$$

ただし $Z = \{Z_n; n = 1, 2, \ldots\}$ は雑音を表す退化していない平均0のガウス過程、$S = \{S_n; n = 1, 2, \ldots\}$ と $Y = \{Y_n; n = 1, 2, \ldots\}$ はそれぞれ入力信号と出力信号を表す確率過程である。通信路は雑音のかからないフィードバックをもつとする。したがって S_n は送信するメッセージと出力信号 $Y_1, \ldots, Y_{n-1}$ の函数であるとして表される。 レート R, 長さ n の符号語 $x^n(W, Y^{n-1})$, $W \in \{1, \ldots, 2^{nR}\}$ と復号函数 $g_n : \mathbb{R}^n \to \{1, 2, \ldots, 2^{nR}\}$ に対して、誤り確率は

$$Pe^{(n)} = Pr\{g_n(Y^n) \neq W; Y^n = x^n(W, Y^{n-1}) + Z^n\},$$

で定義される。ただし W は $\{1, 2, \ldots, 2^{nR}\}$ 上の一様分布で雑音 $Z^n = (Z_1, Z_2, \ldots, Z_n)$ とは独立である。 入力信号には平均電力制限が課せられる。つまり

$$\frac{1}{n} \sum_{i=1}^{n} E[S_i^2] \le P$$

である。またフィードバックは causal である。つまり $S_i (i=1,2,\ldots,n)$ は $Z_1,\ldots,Z_{i-1}$ に従属している。同様にフィードバックがない場合は S_i $(i = 1,2,\ldots,n)$ は $Z^n = (Z_1, Z_2,\ldots,Z_n)$ と独立である。
有限ブロック長容量を次のように定義する。

$$C_{n,FB}(P) = \max \frac{1}{2n} \log \frac{|(I+B)R_Z^{(n)}(I+B)^t + R_V^{(n)}|}{|R_Z^{(n)}|},$$

ただし最大値は

$$Tr[BR_Z^{(n)}B^t + R_V^{(n)}] \leq nP$$

を満たす狭義下三角行列 B と 非負対称行列 $R_V^{(n)}$ についてとる。同様にフィードバックがないときには容量 $C_n(P)$ は $B=0$ としたときの最大値である。これらの条件の下で Cover and Pombra は次の結果を得た。

Theorem 3 (Cover-Pombra [2]) 任意の $\epsilon > 0$ に対して各 $n = 1,2,\ldots$ でブロック長 n で $2^{n(C_{n,FB}(P)-\epsilon)}$ 個の符号語が存在して $n \to \infty$ のとき $Pe^{(n)} \to 0$ とできる。逆に任意の $\epsilon > 0$ とブロック長 n で $2^{n(C_{n,FB}(P)+\epsilon)}$ 個の符号語からなる任意の符号の列に対しても $Pe^{(n)} \to 0$ $(n \to \infty)$ が成り立たない。これはフィードバックをもたない場合も成り立つ。

ここではブロック長 n を固定したとき $C_{n,FB}(P)$ と $C_n(P)$ との間の関係に興味がある。$C_n(P)$ は正確に求められている。

Proposition 1 (Gallager [5])

$$C_n(P) = \frac{1}{2n}\sum_{i=1}^{k} \log \frac{nP + r_1 + \cdots + r_k}{kr_i},$$

ただし $0 < r_1 \leq r_2 \leq \cdots \leq r_n$ は $R_Z^{(n)}$ の固有値、$k(\leq n)$ は $nP + r_1 + \cdots + r_k > kr_k$ を満たす最大整数である。

ところで $C_{n,FB}(P)$ は正確には得られないので、今まで多くの人々によって様々な形の上界が得られている。例えば Ebert [4], Pinsker [10], Cover-Pombra [2], Dembo [3], Yanagi [13] [14] などがある。ここではある意味で強い上界を求める。

$R_Z^{11}(k)$ を $1,\ldots,k-1$ 行 $1,\ldots,k-1$ 列からなる $R_Z^{(n)}$ の部分行列、$R_Z^{12}(k)$ を $1,\ldots,k-1$ 行、$k,\ldots,n$ 列からなる $R_Z^{(n)}$ の部分行列、$R_Z^{21}(k) = R_Z^{12}(k)^t$ とする. このとき次を得る.

Theorem 4

$$C_n(P) \le C_{n,FB}(P) \le C_n(P^*),$$

ただし

$$P^* = P + \sqrt{\frac{P}{n}}\{\sqrt{\sum_{k=2}^{n} \lambda_{\max}(\begin{pmatrix} R_Z^{11}(k) & R_Z^{12}(k) \\ R_Z^{21}(k) & R_Z^{21}(k)R_Z^{11}(k)^{-1}R_Z^{12}(k) \end{pmatrix})}$$

$$+\sqrt{Tr[R_Z^{(n)}] - |R_Z^{11}(2)| - \frac{|R_Z^{11}(3)|}{|R_Z^{11}(2)|} - \cdots - \frac{|R_Z^{(n)}|}{|R_Z^{11}(n)|}}\}.$$

4 フィードバックをもつ連続時間ガウス型通信路

$(\Omega, \mathcal{F}, \mathcal{P})$ を確率空間とする．雑音 $Z = \{Z(t); 0 \le t \le T(< \infty)\}$ は $(\Omega, \mathcal{F}, \mathcal{P})$ 上で定義された平均 0 の可分なガウス過程で

$$\int_T E[Z(t)^2]dt < \infty$$

を満たすとする．このときフィードバックをもつ不適合型連続時間ガウス型通信路は次のように与えられる．

$$Y(t) = S(t) + Z(t),\ 0 \le t \le T,$$

ただし $S = \{S(t); 0 \le t \le T\}$ と $Y = \{Y(t); 0 \le t \le T\}$ はそれぞれ $(\Omega, \mathcal{F}, \mathcal{P})$ 上で定義された入力信号過程と対応する出力信号過程である．フィードバックをもつことにより $S(t)$ はメッセージ X と出力信号 $Y(u), 0 \le u \le t$ のある函数として表される．詳しくは省略する．$\|\cdot\|_W$ を Z とは異なる W の再生核ヒルベルト空間 H_W のノルムであるとすると、S に対する拘束条件

$$E\|S\|_W^2 \le P$$

の下での通信路の容量を考える．$C_0(P)$ をフィードバックをもたない場合の容量、$C_f(P)$ をフィードバックをもつ場合の容量とする．$C_0(P)$ の値は Baker [1] により正確に得られているが、$C_f(P)$ に関しては離散時間のときと同様に全く得られていない．したがって $C_f(P)$ の上界を求めたい．次のような仮定をする必要がある．

(G1) μ_W と μ_Z が strongly equivalent である．即ち

$$R_Z = R_W^{1/2}(I + K)R_Z^{1/2},\ \ K \in (\tau c).$$

ここで (τc) はトレース・クラスの作用素全体を表す．

(G2) X のほとんどのパスが H_W に属する. 即ち

$$\mu_X[H_W] = 1.$$

(G3) $S = X - TY$, ただし $range(T) \subset H_W$, でかつ

$$T \in (\tau c),\ \sigma(T) = \{0\}.$$

ここで $\sigma(T)$ は T のスペクトルを表す.

次の定理を得る.

Theorem 5

$$C_0(P) \leq C_f(P) \leq 2C_0(P).$$

Theorem 6 次の (1), (2) が成り立つ.

(1) $C_0(P) \leq C_f(P) \leq \dfrac{1}{2}||(I+K)^{-1}||P,$

(2) $K = 0$ 又は $K \geq 0$ のとき $C_0(P) = C_f(P) = \dfrac{P}{2}.$

Theorem 7 次の (1), (2), (3) が成り立つ.

(1) $C_0(P) \leq C_f(P) \leq C_0(P^*)$,
ただし

$$P^* = P + 2\sqrt{P}\sqrt{\sigma[K(I+K)^{-1/2}]}$$

である.

(2) すべての $\alpha\ (> \frac{1}{2})$ に対して

$$\lim_{P\to\infty} \frac{C_f(P) - C_0(P)}{P^\alpha} = 0.$$

(3) $\displaystyle\lim_{P\to\infty} \frac{C_f(P)}{C_0(P)} = 1.$

参考文献

[1] C. R. Baker, "Capacity of the mismathched Gaussian channel", IEEE Trans. Information Theory, Vol IT-33, pp 802-812, 1987.

[2] T. Cover and S. Pombra, "Gaussian feedback capacity", IEEE Trans. Information Theory, Vol IT-35, pp 37-43, January 1989.

[3] A. Dembo, "On Gaussian feedback capacity", IEEE Trans. Information Theory, Vol IT-35, pp 1072-1089, September 1989.

[4] P. Ebert, "The capacity of the Gaussian channel with feedback", Bell. Syst. Tech. J., pp 1705-1712, 1970.

[5] R. G. Gallager, Information theory and reliable communication, John Wiley and Sons, New York, 1968.

[6] S. Ihara, "On the capacity of the discrete time Gaussian channel with feedback", Trans. Eighth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, Vol C, Czecho. Acad. Sci., pp 175-186, 1978.

[7] S. Ihara and K.Yanagi, "Capacity of discrete time Gaussian channel with and without feedback, II", Japan J. Appl. Math., Vol 6, pp 245-258, 1989.

[8] 国澤清典、梅垣壽春, 情報理論の進歩, 岩波書店, 1965.

[9] L. Ozarow, "Upper bounds on the capacity of Gaussian channels with feedback", IEEE Trans. Information Theory, Vol IT-36, pp 156-161, January 1990.

[10] M. Pinsker, talk delivered at the Soviet Information Theory Meeting, (no abstract published), 1969.

[11] R. Schatten, Norm ideals of completely continuous operators, Springer-Verlag, Berlin/New York, 1960.

[12] N. N. Vakhania, V. I. Tarieladze and S. A. Chobanyan, Probability distribution on Banach spaces, D.Reidel Publishing Company, 1987.

[13] K. Yanagi, "An upper bound to the capacity of discrete time Gaussian channel with feedback", Lecture Notes in Math., Vol 1299, pp 565-570, 1988.

[14] K. Yanagi, "Necessary and sufficient condition for capacity of the discrete time Gaussian channel to be increased by feedback", IEEE Trans. Information Theory, Vol IT-38, pp 1788-1791, Nobember 1992.

[15] K. Yanagi, "An upper bound to the capacity of discrete time Gaussian channel with feedback, II", IEEE Trans. Information Theory, Vol IT-40, pp 588-593, March 1994.

755 山口県宇部市常盤台２５５７
山口大学工学部共通講座
0836-31-5100 (Ext 4911)
0836-35-9966 (Dial-in)
yanagi@hakucho.apsci.yamaguchi-u.ac.jp (e-mail)